# Managing Internet Surveillance Systems: When the Paper Trail is Gone

Todd Gerber, M.S.

F. Bruce Coles, D.O.

New York State Department of Health



# Objectives

- Review quality control issues
- Understand methods for conducting quality control activities on normalized database structures
- Identify methods to electronically identify errors
- Describe tools for reporting quality control issues to registrars



# Why Quality Data?

- Florida Presidential Vote 2000
- Firestone Tires
- Logan Airport and Massport



## Quality Data is Useful

• Decision support

• A "single version of the truth"

Knowing the clients



## **Quality Control Failures**

- Imagine a mother giving birth to a deformed child from congenital syphilis
- Or a teenager suffering from Pelvic Inflammatory Disease
- Or a young woman who is unable to conceive
- Or uncontrolled syphilis spread



### The Demons of Data

Latency

Lack of relevant data

Inability to analyze

Ineffective sharing





## Dirty Data Ruins Relationships

- Bad operational decisions
- Costly to fix
- Inaccurate views of field productivity, staff effectiveness and disease impact



## **Business Process Changes**

- Paper-based system 1940? 1997
- Paper/Web-based 1997-2002
  - Health Information Network (HIN)
- Web-based 2003-present
- Chlamydia reporting, 2000

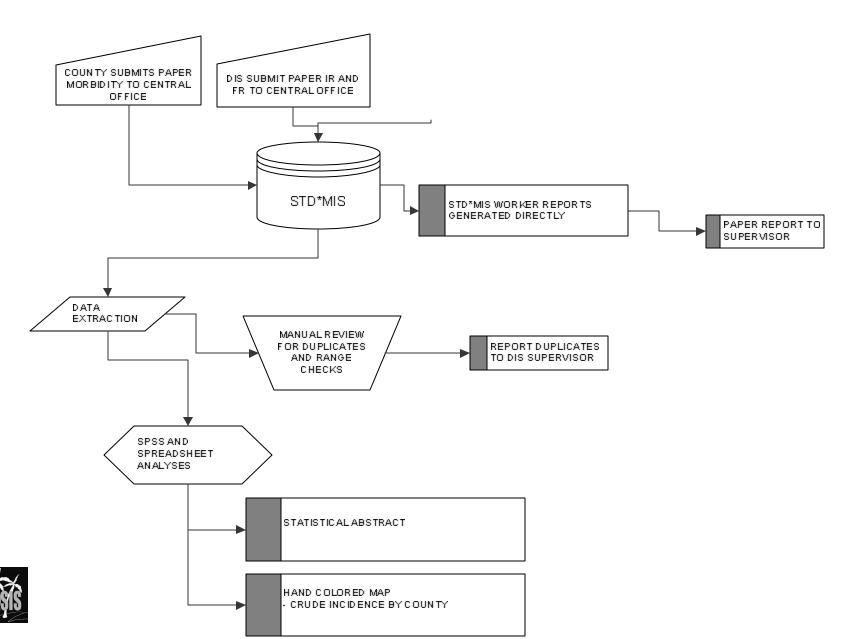


# **Business Process Changes Add Value, Not Simplicity**

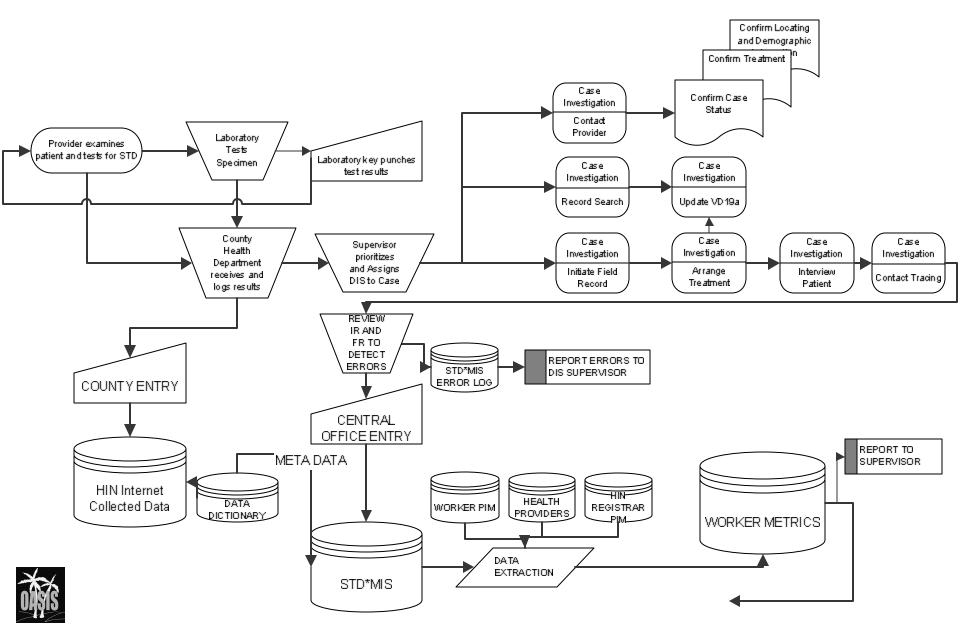
Data Elements		Source	Storage	Entry	Access to Database
1962	Eight	CCR (DC103)	DOH	Batch Unit	NYSDOH
1973	Nine		Mainframe	Punch/Tape Stat Unit	Stat Unit BSTDC
1986	Thirty- seven	CCR (DOH 389)		Stat Unit Epilnfo (4)	
1997	Fifty-six +	CCR or Local Source	Sybase RDBMS HIN ('92-'98 Migrated)	WEB HIN (10 Direct, 2 Upload)	LHU NYSDOH
2005				57 Direct WEB entry	LHU NYSDOH



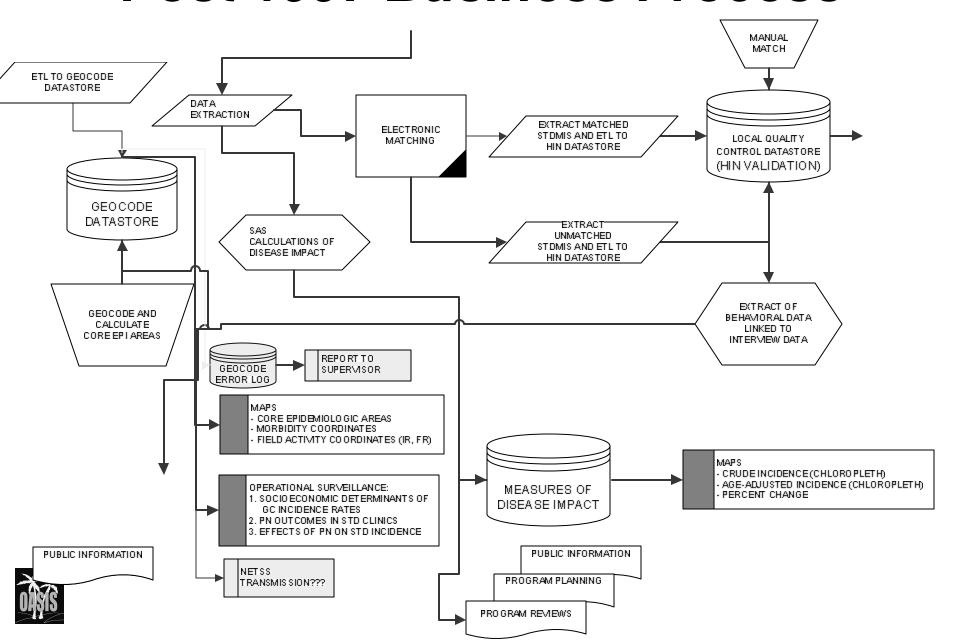
### Pre 1997 Business Process



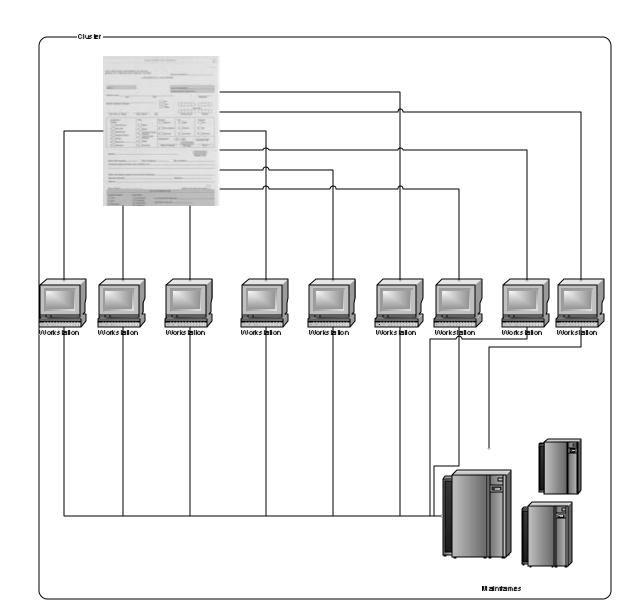
### Post 1997 Business Process



### Post 1997 Business Process

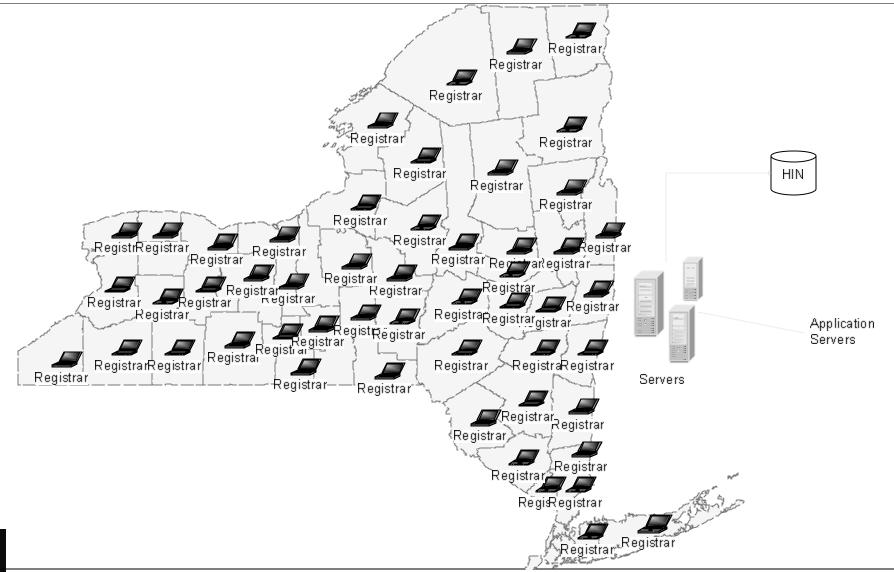


## Pre 1997, Central Data Entry





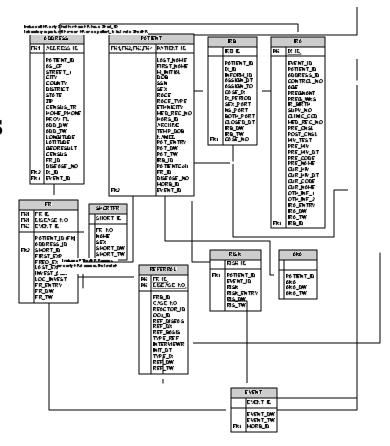
### Post 1997, Real Time Data Entry





### Real-time and Normalized

- Characterized by
  - Relational tables
  - Primary and foreign keys
  - Entity relationships
  - Reduced redundancy
- Performance issues
  - Updates are easy
  - Queries are slow

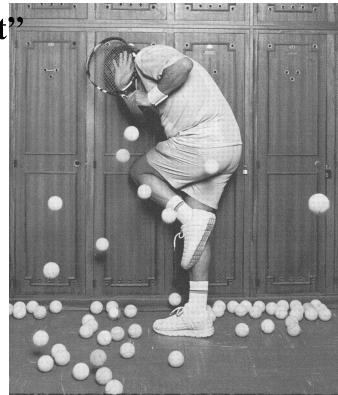




### Real Time, Point-of-Source Data

#### Sometimes, real fast is not good enough

- Real-time availability tempered by latency during data entry
- Constant volume creates "firehose effect"
- Multiple reporters creates feedback problems for errors
- Staff turnover compounds difficulty in tracking and correcting errors

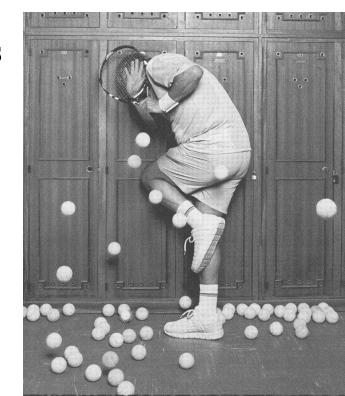




### Real Time, Point-of-Source Data

#### Sometimes, real fast is not good enough

- Data arrives duplicated, missing values, wrong values or inconsistent with case management data
- Unlinked to case management activities and unseen by Disease Intervention Specialists
- Hardcopy "paper trail" to review and compare not available





# Major Issues

- Process errors
- Individual errors
- Nonstandard data
- Nonprecise business rules
- Missing data (incomplete)
- Inconsistencies in reporting
- Timeliness



# Examples

- Errors in recording data (inaccuracy)
- Wrong data or Non-current (addresses)
- Logic errors
  - "Brought to Rx" before date of initiation
  - Male with cervix as site
  - Male pregnant
  - Male names with female sex



#### Nonstandard Results Error

- RPR
- RPR TITER
- FTA-ABS
- T pallidum Ab (FTA-Ab)
- FTA-ABS,SERUM
- RPR W/TITER & CONF RFX
- Syphilis IgG Ab
- TPPA (MHATP)
- FTA
- TPPA
- RPR Titer
- T pallidum Antibodies (TP
- RPR (MONITOR) W/RFX TITER
- RPR,FTA-ABS
- Other Syphilis
- RPR, Rfx Qn RPR/Confirm T
- RPR, TITER
- VDRL (SPINAL FLUID)
- TP-PA
- RPR+Rub Ab+VarAb+RubeoAb+
- RPR/FTA-ABS

- RPR, RFX QN RPR/CONFIRM T
- Rapid Plasma Reagin, Quan
- CBC/D/Plt+RPR+Rh+ABO+Rub
- FTA-ABS (T PALLIDUM) CONF
- Rpr and Mha-tp
- CMP12+LP+6AC+CBC/D/Plt+RP
- Prenatal Profile I
- rpr
- CMP12+8AC+RPR+RA Qn
- Uric A+RPR+ESR-Wes+ANA+RA
- CMP12+LP+TP+TSH+6AC+CBC/D
- CMP12+LP+TP+6AC+CBC/D/Plt
- RPR+Chlamy+HSV1/2
- RPR+Rh+ABO+Rub Ab+Ab Scr+
- rpr and Mha-tp
- CBC/D/Plt+RPR+Gest Glu
- Gc probe, non-amplified
- MHA-TP
- Prenatal Profile IV
- TP-IgG
- VDRL
- CBC/D/Plt+RPR+UA+Rh+ABO+R
- Gc Culture
- Glu+CBC/D/Plt+RPR+UA+Rh+A
- PRENATAL PROFILE I
- RPR SERUM
- Rpr AND Mha-tp
- VDRL, CSF
- rpr and mha-tp
- BUN+Glu+CBC/D/Plt+RPR+UA+



### Reporters = 25,879

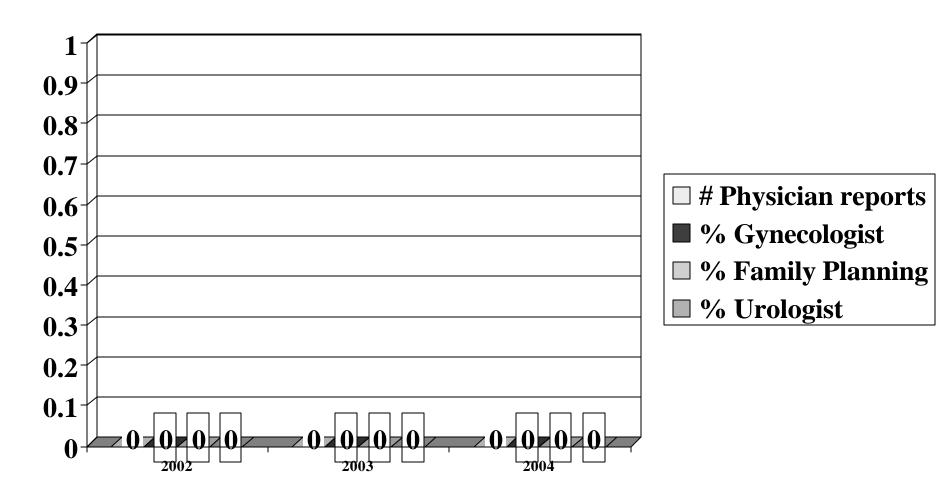
- VASAR BROS ER
- VASSAR BROS CARE CTR
- VASSAR BROS ER
- VASSAR BROS HOSP
- VASSAR BROS HOSP ER
- VASSAR BROS HOSP LABOR & DELIVERY
- VASSAR BROS HOSP/DR YOON
- VASSAR BROS HOSPITAL
- VASSAR BROS HOSPITAL ER
- VASSAR BROS HOSPITAL/ER
- VASSAR BROS HOSPITAL/OB/GYN
- VASSAR BROS, ER
- VASSAR BROS, GYN CLINIC
- VASSAR BROS, HOSPITAL
- VASSAR BROS, HOSPITAL CLINIC
- VASSAR BROS, HOSPITAL ER
- VASSAR BROTHERS
- VASSAR BROTHERS CARE
- VASSAR BROTHERS CLINIC
- VASSAR BROTHERS HOSPITAL
- VASSAR BROTHERS HOSPITAL ER
- VASSAR BROTHERS HOSPITAL OB CLINIC
- VASSAR CAR
- VASSAR CARE
- VASSAR CARE CENTER
- VASSAR CARE CENTER/DR HABIB
- VASSAR CARE CLINIC
- VASSAR CARE CNTR
- VASSAR CARE CTR
- VASSAR CLINIC
- VASSAR CLINIC YOUNG
- VASSAR COLLEGE
- VASSAR COLLEGE HEALTH SERVICE
- VASSAR COLLEGE HEALTH SERVICES
- VASSAR EMERGENCY ROOM
- VASSAR ER
- VASSAR FAST TRACK
- VASSAR GYN CLINIC
- VASSAR HOSPITAL
  - VASSAR HOSPITAL CLINIC
  - VASSAR HOSPITAL ER
  - VASSAR OB
  - VASSAR OB CLINIC
- VASSAR OB/GYN
- VASSSAR BROTHERS ER

Dutchess County HD - STD Clini	34	
Vassar Clinic/ Vassar Brothers	162	
Childrens Medical Group/Adoles	1	
Medicus- Wappingers Falls	1	
Misc PMDs	4	
Vassar College / Maplebrook Sc	4	
Planned Parenthood- Poughkeeps	1	
Vassar Bros Hosp (VBH)-Poughke	100	



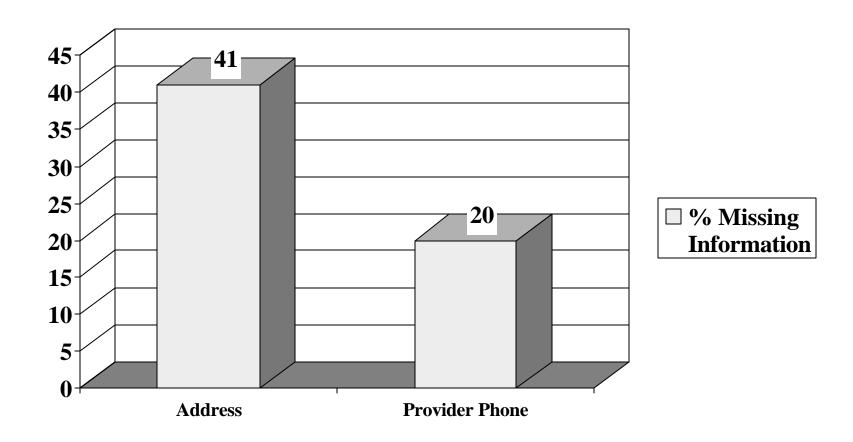
### Physician Specialty by Year of Report

(need to collect data to present data)



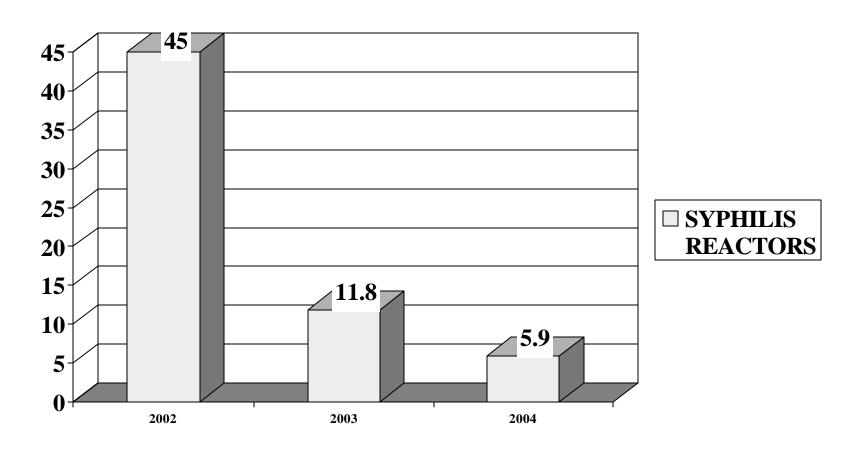


## ECLRS Missing Information



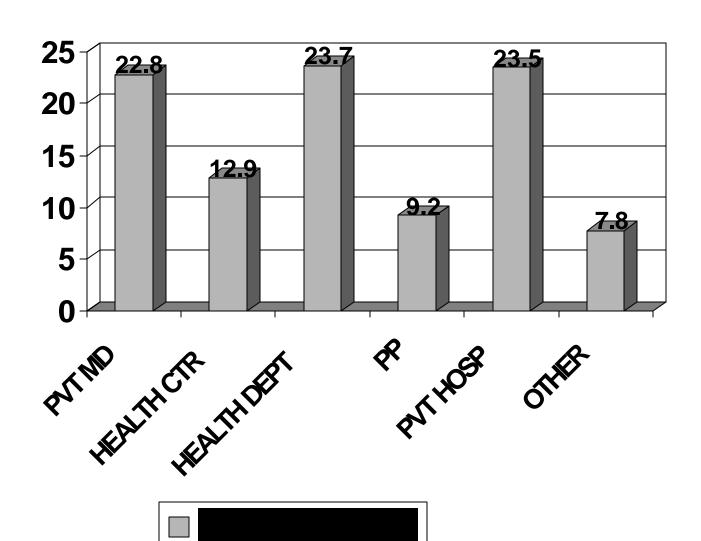


## ECLRS Days to Open



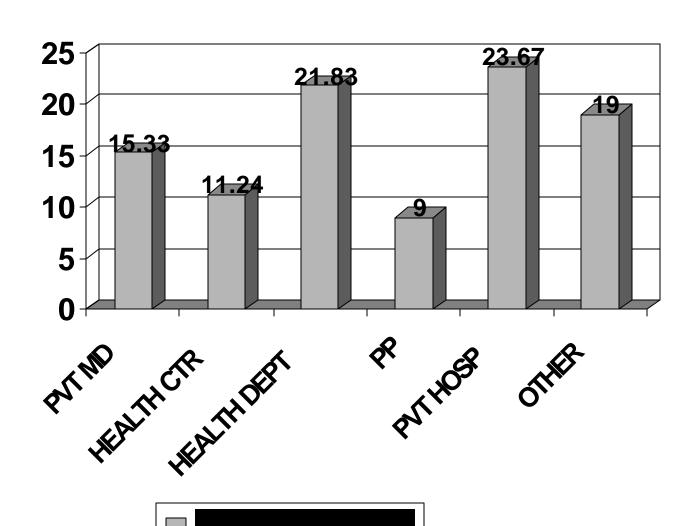


#### Source Code Field





### Source Provider Type





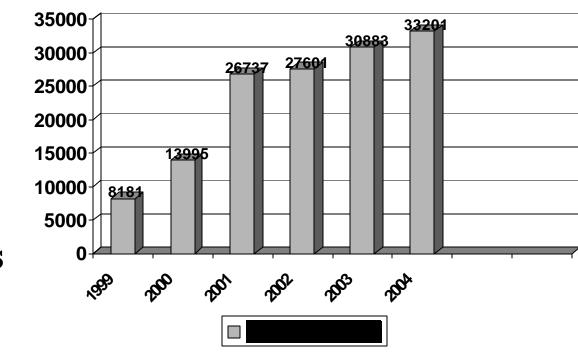
## **Achieving Quality Data**

- Apply standards to produce consistent results
  - Process
  - Data definition
  - Technology
- Design quality into the process and train to minimize need for inspection and interpretation
- Audit and improve existing data



### Record-level Audit?

- Volume may be too large
- Random or structured sampling audit of priority cases
- Audit to ID problems
- Audit to improve training





#### Record-level Audit

- Reconstituted electronic records
- Problem logging to record findings
- Error report joined to automated error detection



## **Process Improvement**

- Standardize data elements
- Improve data collection with better instructions/manuals
- Review errors with data Registrars
- Prepare lookup and domain tables
- Identify and describe table relationships, constraints and dependencies
- Develop queries to de-normalize tables for analysis



### **Four Solutions**

- Recursive queries
- Reference tables
- Electronic logs for recording feedback

Continued training support



### Recursive Query Reference Table



Firm Result field is used to populate the company address dictionary reference table

ADDRESS			
PK FK1 PK	A DORES S_PROBLEM_ID A DORES S_ID		
	PATIENT_ID AS_OF STREET_1 CITY COUNTY DISTRICT STATE ZIP CENSUS_TR HOME_PHONE ADD_DW ADD_TW company AddressType FIPChty Address Notes		

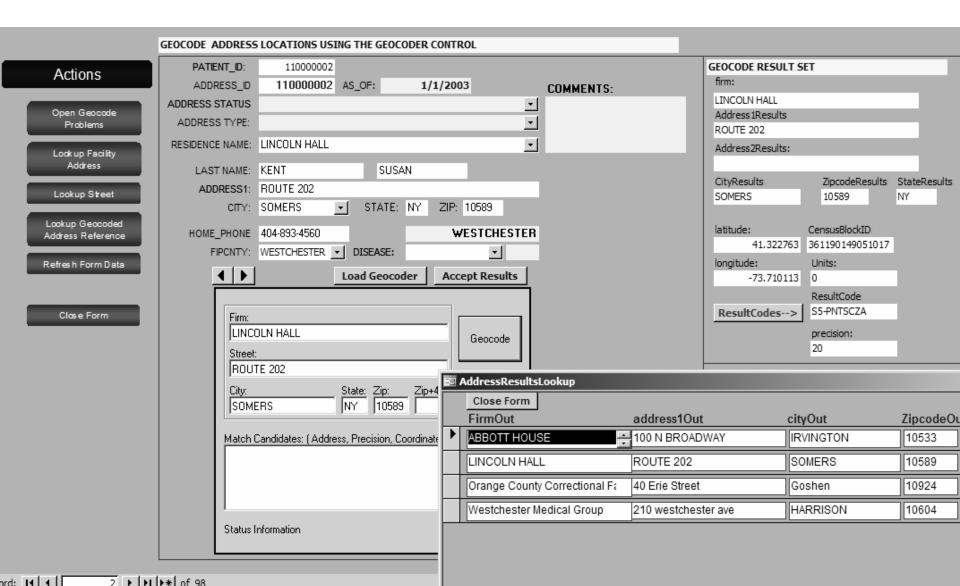
Address Problems					
PK					
II					

Census Block ID Results ADDRESS\_ID PATIENT ID AsOtRes I Is GEORESULT tattede longitude ResunCode Firm Results CompanyResults AddressStatus Address1Results Address2Res | Its CityRes (its County Results District Results State Results Zipcode Results Hom e Phone Results Pie obio a Results DeliueryPointResuits Urbanization Results UnitsResults Comments Results ADD DOV ADD\_TW

Add ess Result



## Recursive Query Searches



# Reference Tables

- Providers (Health Services File)
- Disease Intervention Specialists
- Licensure files
- Streets
- Zip codes
- Facilities
- Communication Directory



## Legacy Reference Table

	<b>?</b>
	015
	017
CATTARAUGUS CTY PHN	018
NIAGARA COUNTY PHNS	019
MONTALVO C,(020)	020
SCHNURR J,(021)	021
FORT DRUM,(022)	022
HENDERSON T,(023)	023
ROSCOE M,(024)	024
SHIPMON C,(025)	025
ETHIER K,(026)	026
BENDER M,(027)	027
FRIEDMAN A,(028)	028
TATE A,(029)	029
CORTI J,(030)	030
STERN A,(031)	031
ROBKOFF R,(032)	032
AVIMELEH J,(033)	033
MOORE S,(034)	034
POLVERARI D,(036)	036
OKPAN C,(037)	037
EVANS A,(038)	038
NEMES E,(039)	039
NALLY G,(041)	041
PARLMAN D,(042)	042
SINGER L,(043)	043
LEVINE J,(044)	044
BARRETT A,(045)	045
ZIMBLER C,(046)	046
KULESA E,(047)	047
LUGO N,(048)	048



### RDBMS Reference Table

WORKERI	LNAME	FNAME	REGION	COUNTY	STATUS	COMPANY	DEPART
	Ayers	Lisa	ALBANY	SCHENECTADY	CURRENT	Schenectady County DOH	
586	BARKER	LENNOX	MARO	NASSAU	current	NASSAU COUNTY HEALTH DE	BUREAU OF STD CONT
564	BARRETO	DORA	MARO	Dutchess	CURRENT	COUNTY HEALTH DEPARTME	BUREAU OF STD CONT
562	BARRETT	PAUL	MARO	Dutchess	CURRENT	DUTCHESS COUNTY DEPT OF	BUREAU OF STD CONT
572	BAYLES	MICHAEL	MARO	MILITARY	MILITARY	WEST POINT	KELLER ARMY COMMI
361	BECRAFT	REBECCA	ROCHESTER	CHEMUNG	CURRENT	CHEMUNG COUNTY HEALTH	
538	BERNARD	RALPH	MARO	Westchester	CURRENT	WESTCHESTER COUNTY DEF	STD CONTROL

current

MILITARY

CURRENT

CURRENT

MILITARY

current

current

current

current

current

ONEIDA

Suffolk

ONONDAGA

DUTCHESS

Westchester

Westchester

ONONDAGA

JEFFERSON

SUFFOLK

Nassau

Dutchess

DUTCHESS

DUTCHESS

**JEFFERSON** 

**JEFFERSON** 

MILITARY

REGIONAL OFFICE

REGIONAL OFFICE

REGIONAL OFFICE

REGIONAL OFFICE

REGIONAL OFFICE

REGIONAL OFFICE

WESTCHESTER

PENNY

EVE

JUDI

**CHRIS** 

Massie

**RALPH** 

ENRICO

DEBORA

CORLETA

**JESSICA** 

DIERDRE

MICHELE

ANDREW

CHARLENE

LARRY

ANNE

PAT

SUSAN

NANCY

BRIAN

**PATRICIA** 

MARY PAT

KIMBERLY

SYRACUSE

SYRACUSE

SYRACUSE

MARO

MARO

MARO

MARO

ALBANY

ALBANY

SYRACUSE

ROCHESTER

MARO

MARO

MARO

MARO

MARO

MARO

MARO

MARO

**BUFFALO** 

SYRACUSE

SYRACUSE

SYRACUSE

BLASZKOW

BOWERS

BRANDON

BROWN

BRYANT

BULMER

BURNETT

CAPRARI

CROUCH

CUSANO

DANIELS

DOBIES

**EVANS** 

**EVANS** 

FIXX

DONOHUE

EBERLING

FRIEDMAN

GEAGAN

HAITH

CUMMINGS

Brown

BOYLE

458

527

561

403

585

576

110

106

504

411

303

542

255

456

511

565

506

524

455

520

457

571

NYS DEPARTMENT OF HEAL'

NYS DEPARTMENT OF HEAL

ERIE COUNTY DEPARTMENT

FORT DRUM

WEST POINT

ONEIDA COUNTY HEALTH DE 520 SENECA ST

NYS DEPARTMENT OF HEAL BUREAU OF DISEASE

SUFFOLK COUNTY DEPARTM SUFFOLK COUNTY STI

DUTCHESS COUNTY DEPT OF COMMUNICABLE DISE

Westchester County Departmel Communicable Disease

NEW YORK STATE DEPARTM ALBANY REGIONAL OF

ONONDAGA COUNTY HEALTH BUREAU OF DISEASE

NYS DEPARTMENT OF HEAL ROCHESTER REGION

COUNTY HEALTH DEPARTME BUREAU OF STD CON

DUTCHESS COUNTY DEPART COMMUNICABLE DISE

PREVENTATIVE MEDIC

SUFFOLK COUNTY DEPARTM SUFFOLK CTY STD

WESTCHESTER COUNTY DEFISTD CONTROL

WESTCHESTER COUNTY DEF STD CONTROL

NASSAU COUNTY DEPARTME STD CONTROL

DUTCHESS COUNTY DEPART COMMUNICABLE DISE CURRENT MILITARY PREVENTATIVE MEDICINE SEATTN COMMUNITY HE CURRENT DEPARTMENT OF HEALTH COMMUNICABLE DISE PREVENTATIVE MEDIC MILITARY FORT DRUM

## Legacy SAS Format

```
*/
* last update 12-10-01;
* stdmis worker codes;
value $worker
  '000'
            'UNKNOWN
  '015'
            'CAPRARI EIIII
  '017' =
            'CHAUTAUOUA CTY PHNS
 '018' =
            'CATTARAUGUS CTY PHNS
  '019' =
            'NIAGARA COUNTY PHNS
  '020' =
            'MONTALVO C
  '021' = 'SCHNURR J
  '022' = 'FORT DRUM
  '023' = 'HENDERSON T
  '024' =
            'ROSCOE M
  '027' =
            'BENDER M
  '028' =
            'FRIEDMAN A
  '029' =
            'TATE A
  '030' =
            'CORTI J
  '031' =
            'STERN A
  '032' =
            'ROBKOFF R
  '033'
            'AVIMELEH J
  '034' =
            'MOORE S
  '035' =
            'DEBRA GRACIA NASSAU COUNTY
  '036'
            'POLVERARI D
  '037'
            'OKPAN C
  '038' =
            'EVANS A
  '039' =
            'NEMES E
  '040'
            'SUE PAYETTE
 '041'
            'NALLY G
 '042' =
            'PARLMAN D
  '043' =
            'SINGER L
  '044'
            'LEVINE J
```



====>

### RDBMS SAS Format

```
* CREATE WORKER NUMBERFORMAT using the Access Query named grySASUserFormat;
/* PROC IMPORT OUT= work.workernumberformat
            DATATABLE= "grySASformatWorkerno"
            DBMS=ACCESS REPLACE:
      DATABASE="H:\DOCS\DOCUMENT FOLDERS\ACCESS DATA\STD CONTACTS\STD CONTACTS 2000.mdb";
     USERID="todd":
     PASSWORD="&ACCESSPW2";
     WORKGPDB="H:\DOCS\HIN VALIDATION\MARK.mdw":
                                                            ———OPEN THE DATABASE
         RUN: */
proc sql :
   connect to sqlservr (server=SDOHCCH22\SQLSTDMIS database="BSTDC MANAGERSQL");
CREATE table WORK workernumberformat as SELECT *
                        FROM CONNECTION TO sqlservr (SELECT * FROM grySASformatWorkerno);
options nodate pageno=1 linesize=80 pagesize=60;
  data ctrl(rename=(WORKERNO=start name=label));
   length WORKERNO $3. Iname $35.;
                                                         Read the Query as a SAS Format
   set work.workernumberformat :
   retain fmtname 'worker' type 'c':
run;
     proc print data=ctrl noobs:
         var start label:
      title 'The workerno Data Set':
   run;
        proc format library=library cntlin=ctrl;
   run;
   quit;
     proc print data=workerno noobs:
         var start label;
```

title 'The workerno Data Set':

## Problem Logging for Feedback

#### ADDRESS PROBLEMS FORM



PATIENT_ID: 110000001	_	RETURN STATUS: NOT RETURNED	WORKER: ▼ BOMEISL, KE	NNETH, 501					
Iname: BANDA	fname: ENRIQUE	GeoResult:	DATE_RETUR	RNED MA					
address: 104 FULTON AVE.	city: WHITE F	PLAINS	zip: 12603	COUNTY: WESTCHESTER					
GEOCODE ERROR: DIRECTIONAL (E	, W, N, S) NOT SPEC	IFIED		•					
GEOCODE STATUS MANUAL GEOCO	- ENTER REASON IF N	NONGEOCODABLE:		•					
ERROR DESCRIPTION: This street address requires a directional such as north, south, east, west. Please indicate the directional and return the report to the Bureau of STD Control.									
Creator of Record	TMG02	Edited By:	TMG02						
Date Entered	4/14/2005 10:20:2		5/7/2005 3:09	9:36 PM					



# Report to Registrar

#### HIN 2004 REPORTS THAT DO NOT GEOCODE

PLEASE REVIEW THE ATTACHED HIN REPORTS. THE HIN REPORTS LISTED WERE NOT GEOCODABLE. A GEOCODE ERROR IS PROVIDED THAT DESCRIBES THE REASON THAT THIS REPORT WAS NOT GEOCODED. PLEASE CALL JANELLE ARM STRONG-BROWN TO REVIEW AND CORRECT THESE REPORTS.

RESPONSES MAY BE REVIEWED WITH:

JANELLE ARMSTRONG-BROVIN (518) 474-3598

WESTCH	IESTER	COUNTY HIN	2004 REPORTS	THAT DO	NOT GEOCODE				
address_i ਕ	last name	first name mi	street address	zip	city	geocode error	error description	creator	date entered
WESTCHES	STER C	OUNTY							
110000006	ALTMAN	RON	14 GREENDALE RD	10583	SCARSDALE	OTHER ERROR (SEE COMMENTS)	ADDRESS TYPE USTED AS APARTMENT COMPLEX, IS THIS CORRECT?	M DORE, SEAN, 100	4/28/2004
110000033	BAKER	RICK		10606	WHITE PLAINS	BLANK STREET AD DRESS LISTED	glesse complete the street address field or identify the address for the grounder of this region and indicate that this should be geolooded to the	GERBER, 7000, 200	10/9/2004
110000001	BANDA	ENRIQUE	104 FULT ON AVE.	12603	WHITE PLAINS	DIRECTIONAL (E, W, N, S) NOT SPECIFIED	This street address requires a directional such as north, south, east, west. Please indicate the directional and return the report to the the Suresu of STO Control.		4/28/2004
110000054	BROOKS	STEVE		10606	WHITE PLAINS	BLANK STREET AD DRESS LISTED	PLE ASE ENTER A STREET ADDRESS	M DORE, SEAN, 100	8/3/3004
110000014	BRUCE	COLES	72 VALLEY ROAD	10583	SCARSDALE	NO SUCH STREET IN THIS CTV		MOORE, SEAN, 100	8/13/2004
110000005	CRAWFORD	MELVINA	44 GREENDALE RD	10583	SCARSDALE	NO SUCH STREET IN THIS CTV		M OORE, SEAN, 100	4/28/2004



## Training Support

- Communication directory of registrars
- Transaction logs
- Training inventory
- Road shows an on-site training
- Guidelines (PDF, e-mail, Web based)
- SAS license
- SAS training

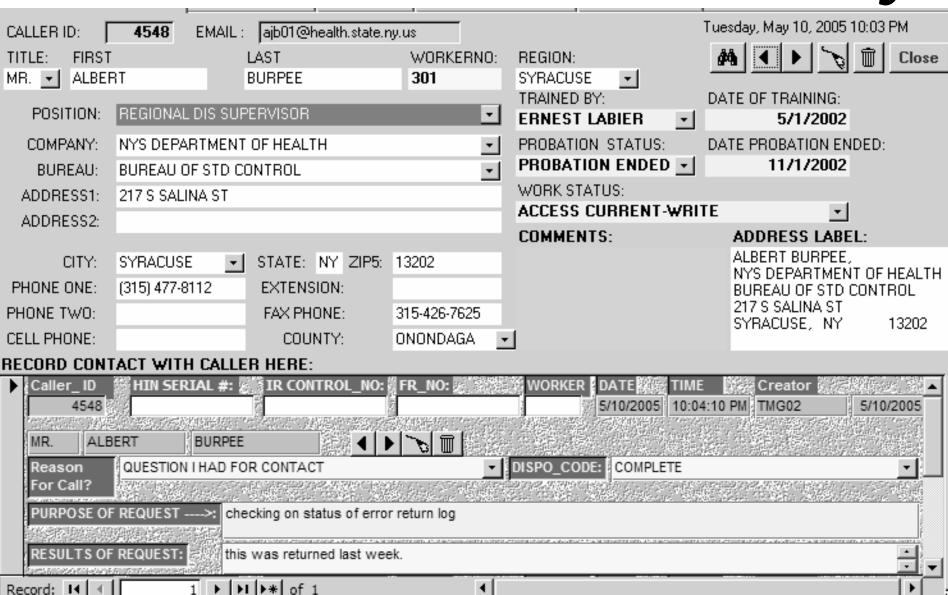


## Communication Directory

- Tracks Registrars
- Records communication details
- Maintains training inventory and status



# **Communication Directory**



# Steps to Improve Quality



## **Quantify the Pain**

- Identify the types of errors that are occurring in your data stream
- Identify the sources of those errors
  - Processes
  - Individuals
- Establish rules for collecting data
- Establish standards for data elements





#### Follow Rules of Thumb

• Meet the most important data needs of the most important customers

Prevention is #1, then cleanup



### Follow Rules of Thumb

Establish processes to prevent errors

Then clean up existing errors

 Start with critical data suppliers-insist they provide accurate and current data



#### Follow Rules of Thumb

• Imbed linking fields for validation against reference tables and external data

Standardize data collection

Establish feedback loop with providers



# Thank You!

Todd Gerber NYS Department of Health tmg02@health.state.ny.us (518) 474-3598

